

## REES MATRIX CONSTRUCTIONS FOR CLUSTERING OF DATA

A. V. KELAREV , P. WATTERS and J. L. YEARWOOD

(Received 7 July 2008; accepted 21 March 2009)

Communicated by M. G. Jackson

### Abstract

This paper continues the investigation of semigroup constructions motivated by applications in data mining. We give a complete description of the error-correcting capabilities of a large family of clusterers based on Rees matrix semigroups well known in semigroup theory. This result strengthens and complements previous formulas recently obtained in the literature. Examples show that our theorems do not generalize to other classes of semigroups.

*2000 Mathematics subject classification:* primary 16S36, 20M35; secondary 20M25, 68T.

*Keywords and phrases:* Rees matrix semigroups, clustering, data mining.

### 1. Introduction

The investigation of semigroup constructions motivated by applications in data mining was begun in [29], where a sophisticated formula was obtained for multiple classifiers based on two-sided ideals in a semigroup construction defined with Brandt semigroups.

The present paper is motivated by applications to the more difficult problem of clustering, and deals with more general constructions. We introduce a new type of multiple clustering systems, or clusterers, based on Rees matrix semigroups, which are well-known technical tools of semigroup theory (see [18]). Let us also refer, for example, to [21, 22, 34] for recent results concerning this construction. The class of all Brandt semigroups is a proper subclass of the class of all Rees matrix semigroups.

Our second improvement is explained by the fact that multiplication in these matrix constructions is not commutative and the family of arbitrary one-sided ideals is much

---

The first author was supported by Discovery Grant DP0449469 from the Australian Research Council. The second author was supported by Linkage Grant LP0776267 from the Australian Research Council. The third author was supported by a Queen Elizabeth II Fellowship and Discovery Grant DP0211866 from the Australian Research Council.

© 2009 Australian Mathematical Publishing Association, Inc. 1446-7887/2009 \$16.00

larger than that of two-sided ideals (see Section 3 for complete definitions). It is essential to consider all ideals not only in order to develop theoretical foundations, but also since the larger set of ideals may lead to the design of clustering systems with better properties. The aim of this paper is to complement and strengthen the results of [29] by handling the case of all one-sided ideals.

Our main theorem gives a formula for the number of errors of individual clusterers which can be corrected by a combined multiple clusterer of this kind (see Theorem 4.1 in Section 4). It follows from a technical result of independent theoretical interest (see Proposition 4.2 in Section 4). The proofs and examples show that our main results do not extend to larger classes of semigroups. An interesting open question is included in Section 7 with a discussion of how our theoretical results of independent interest can be used to guide the design of future experimental investigations.

## 2. Motivation

Classification and clustering of data play central roles in data mining (see, for example [36, 40], and [4, 5, 7, 27, 30, 38]). As a novel example of application let us mention that classification and clustering have recently been used in internet commerce security (see [33]).

Classification deals with known classes of data. These classes are represented by given samples of data. The samples are used for supervised training of the classifier to enable it to recognize new elements of the same known classes. On the other hand, clustering handles data without known sensible groupings or clusters. The task of a clustering system, or clusterer, is to conduct unsupervised investigation of the data in order to determine new groupings or clusters. Although it is possible to optimize and train multiple clusterers in order to increase their ability to handle new data sets in the future, every clustering is undertaken in an unsupervised fashion where the clusters are unknown in advance (see [36, 40] for more details).

A well-known method for designing multiple clusterers consists in designing several binary clusterers (each of which divides the data set into two clusters), and then combining them into one multiple clustering scheme with several clusters. This method is very effective, and is often recommended for various applications (see Witten and Frank [36, Section 7.5]). The main advantage of using combined multiple clusterers is in their ability to correct errors of individual binary clusterers and produce correct clusterings despite individual clustering errors. It is usually desirable to choose a convenient representation for the cluster set of the multiple clusterer and to ensure that it has a small set of generators.

Denote the binary clusterers being combined by  $b_1, \dots, b_m$ . Each of these clusterers divides its input data into two clusters by producing an output 0 or 1 for each input element. If  $o_1, \dots, o_m$  are the outputs of the binary clusterers, then the sequence  $(o_1, \dots, o_m)$  is called a *cluster vector* of the combined multiple clusterer, and the set of all cluster vectors is called the *cluster set*. Each cluster vector represents one cluster in the clustering produced by the multiple clusterer (see [36, Table 7.1]). Every clusterer is uniquely determined by its cluster set.

Let us review the basic essential properties required of the cluster sets. Denote by  $\mathbb{F} = \text{GF}(2)$  the finite field of order 2, that is, the set  $\{0, 1\}$  with the standard addition and multiplication. Denote by  $\mathbb{F}^m$  the set of all sequences of all possible outcomes of the binary clusterers which can occur in general. This means that

$$\mathbb{F}^m = \{(r_1, \dots, r_m) \mid r_1, \dots, r_m \in \mathbb{F}\}.$$

For every element  $r \in \mathbb{F}^m$ , denote by  $r_i$  the  $i$ th component of the sequence  $r$ , so that  $r = (r_1, \dots, r_m) \in \mathbb{F}^m$ .

The *weight* of a sequence  $c$  in  $\mathbb{F}^m$  is denoted by  $\text{wt}(c)$ . It is defined as the number of nonzero coordinates in the vector  $c$ . The *weight* of a cluster set  $C$  is the minimum weight of a nonzero element in  $C$ . The *information rate* of a cluster set  $C$  in  $\mathbb{F}^m$  is defined as the number  $\log_2(|C|)/m$ .

The *minimum distance* of a cluster set  $C$  is the minimum weight among all weights of nonzero differences between pairs of elements in  $C$ . If the cluster set  $C$  forms a linear subspace of  $\mathbb{F}^m$ , then it is very well known and easy to verify that its minimum distance is equal to its weight.

For any real number  $x$ , denote by  $\lfloor x \rfloor$  the *integral part* of  $x$ , or the *floor* of  $x$ , that is the largest integer which does not exceed  $x$ . It is well known and easy to verify that the number of errors of binary clusterers, which the multiple clusterer can correct, is equal to  $\lfloor (d - 1)/2 \rfloor$ , where  $d$  is the minimum distance of the cluster set of the clusterer.

All sequences of the cluster set  $C$  can be written down in a matrix  $M_C$  to discuss their properties. If  $M_C$  has two identical columns, this means that two binary clusterers produce identical outputs. This duplication is very inefficient, even though it could help to correct clustering errors. Therefore, in a situation like this, one of these clusterers can be removed and a better scheme can be devised. Likewise, it is undesirable to have strong correlation or functional dependencies between very small sets of columns in  $M_C$  or between binary clusterers.

According to [36, Section 7.5], for a clusterer with a cluster set  $C$  to be efficient, the cluster set  $C$  must satisfy the following most essential basic properties.

- (A1) The minimum distance of  $C$  must be large.
- (A2) The information rate of  $C$  must be large.
- (A3) A convenient method of generating the set  $C$  is essential.
- (A4) If all vectors of  $C$  are recorded in a matrix  $M_C$ , then there should not be strong correlation or functional dependencies between small sets of columns of  $M_C$ . In particular, the matrix  $M_C$  should not have duplicate columns.

Additional properties may be required depending on the practical application being considered (see Section 7).

### 3. Preliminaries

For the convenience of readers we include some concise preliminary observations and definitions on semigroups. We use standard concepts and refer readers to [18, 24, 26, 36, 40], for more information on semigroup theory, data mining and clustering (see also [1, 6–9, 14, 16, 19, 25, 28]).

It is convenient to generate a cluster set  $C$  with a small number of generators, instead of storing the whole large set  $C$  in computer memory. To this end we introduce addition and multiplication on the set  $\mathbb{F}^m$ . This will enable multiplication between generators and arbitrary elements of  $\mathbb{F}^m$  and the generation of sums of these products.

As usual the standard addition is defined on  $\mathbb{F}^m$  componentwise, that is, the sum of two arbitrary sequences  $(r_1, \dots, r_m)$  and  $(s_1, \dots, s_m)$  in  $\mathbb{F}^m$  is defined as

$$(r_1, \dots, r_m) + (s_1, \dots, s_m) = (r_1 + s_1, \dots, r_m + s_m).$$

In order to generate clusterers with known properties and find optimal multiple clustering schemes, we take a finite semigroup  $S$  and use its ring to introduce additional structure on the cluster set of a multiple clusterer. If  $S$  has a zero, then it will be denoted by  $\theta$ . The number of nonzero elements in  $S$  will be equal to the number of binary clusterers being combined. In other words, we assume that

$$S \setminus \theta = \{s_1, \dots, s_m\}.$$

The *semigroup ring*  $\mathbb{F}[S]$  is the set

$$\mathbb{F}[S] = \left\{ \sum_{s \in S} f_s s \mid f_s \in \mathbb{F} \right\}$$

with addition and multiplication defined by the associative and distributive laws and the rules

$$\begin{aligned} \sum_{s \in S} f_s s + \sum_{s \in S} f'_s s &= \sum_{s \in S} (f_s + f'_s) s, \\ \left( \sum_{s \in S} r_s s \right) \left( \sum_{t \in S} r'_t t \right) &= \sum_{s, t \in S} (r_s r'_t) st. \end{aligned}$$

If  $S$  has a zero  $\theta$ , then a *contracted* semigroup ring is denoted by  $\mathbb{F}_0[S]$  and is defined as the quotient ring of  $\mathbb{F}[S]$  modulo the ideal  $\mathbb{F}\theta$ . Notice that if  $S$  has no zero, then  $S^0$  stands for the semigroup  $S \cup \{\theta\}$  with zero  $\theta$  adjoined; and  $\mathbb{F}[S]$  is isomorphic to  $\mathbb{F}_0[S^0]$ . If  $S$  is a semigroup without zero, then we also let  $\mathbb{F}_0[S] = \mathbb{F}_0[S^0] \cong \mathbb{F}[S]$ . The present paper uses contracted semigroup rings, which helps to record our results more concisely. These constructions are used and considered, for example, in [1, 9, 11, 12, 14, 16, 32].

Regarded as an abelian group, the set  $\mathbb{F}^m$  is isomorphic to the additive group of the contracted semigroup ring  $\mathbb{F}_0[S]$ . It is natural to regard  $S$  as being embedded in  $\mathbb{F}_0[S]$  by identifying each element  $s$  of  $S$  with  $1s$  in  $\mathbb{F}_0[S]$ .

In order to introduce an additional operation on the cluster set  $\mathbb{F}^m$ , we identify the set  $\mathbb{F}^m$  with the contracted semigroup ring  $\mathbb{F}_0[S]$  by identifying each sequence  $r \in \mathbb{F}^m$  with the element  $\sum_{i=1}^m r_i s_i$  of the contracted semigroup ring  $\mathbb{F}_0[S]$ . This means that

$$(r_1, \dots, r_m) = r_1 s_1 + \dots + r_m s_m \in \mathbb{F}_0[S] = \mathbb{F}^m.$$

Given an arbitrary element  $r = r_1s_1 + \dots + r_ms_m \in \mathbb{F}_0[S]$  and any  $1 \leq i \leq m$ , we introduce the notation  $r_{s_i} = r_i$ . This allows us to rewrite any element  $r \in \mathbb{F}^m = \mathbb{F}_0[S]$  as

$$r = \sum_{s \in S} r_s s.$$

Thus, the set  $\mathbb{F}^m = \mathbb{F}_0[S]$  has been endowed with a product defined, for  $x, y \in \mathbb{F}_0[S]$ , by the rule

$$\sum_{s \in S} x_s s \cdot \sum_{t \in S} y_t t = \sum_{s, t \in S} (x_s y_t \cdot st).$$

Clearly, the *weight*  $\text{wt}(r)$  of an element

$$r = \sum_{s \in S} r_s s \in \mathbb{F}_0[S] = \mathbb{F}^m$$

coincides with the number of nonzero coefficients  $r_s$  in  $r$ .

Every element of the form  $fs$ , where  $f \in \mathbb{F}$  and  $s \in S$ , is called a *homogeneous element* of  $\mathbb{F}_0[S]$ . Each term  $r_s s$  is called a *homogeneous component*, or an *s-component*, of the element  $r = \sum_{s \in S} r_s s$ . The *support* of an element  $r \in \mathbb{F}_0[S]$  is defined as the set

$$\text{supp}(r) = \{s \in S \mid r_s \neq 0\}.$$

The weight of  $r$  is equal to the cardinality of  $\text{supp}(r)$ .

Let  $g_1, \dots, g_k \in \mathbb{F}^m$ . An *ideal* or *two-sided ideal* generated by the elements  $g_1, \dots, g_k$  is the set denoted by  $\text{id}(g_1, \dots, g_k)$  and defined by the equality

$$\begin{aligned} \text{id}(g_1, \dots, g_k) &= \mathbb{F}^m g_1 \mathbb{F}^m + \dots + \mathbb{F}^m g_k \mathbb{F}^m \\ &= \left\{ \sum_{j=1}^{m_1} \ell_{1,j} g_1 r_{1,j} + \dots + \sum_{j=1}^{m_k} \ell_{k,j} g_k r_{k,j} \mid \ell_{i,j}, r_{i,j} \in \mathbb{F}^m \cup \{1\} \right\}. \end{aligned}$$

Two-sided ideals in  $\mathbb{F}^m$  were considered in [29]. The present paper deals with one-sided ideals, that is, left ideals and right ideals. A *right ideal* generated by the elements  $g_1, \dots, g_k$  is the set denoted by  $\text{id}_r(g_1, \dots, g_k)$  and defined by

$$\begin{aligned} \text{id}_r(g_1, \dots, g_k) &= g_1 \mathbb{F}^m + \dots + g_k \mathbb{F}^m \\ &= \{g_1 r_1 + \dots + g_k r_k \mid r_1, \dots, r_k \in \mathbb{F}^m \cup \{1\}\}. \end{aligned}$$

A *left ideal* generated by  $g_1, \dots, g_k$  is the set denoted by  $\text{id}_\ell(g_1, \dots, g_k)$  and defined by

$$\begin{aligned} \text{id}_\ell(g_1, \dots, g_k) &= \mathbb{F}^m g_1 + \dots + \mathbb{F}^m g_k \\ &= \{\ell_1 g_1 + \dots + \ell_k g_k \mid \ell_1, \dots, \ell_k \in \mathbb{F}^m \cup \{1\}\}. \end{aligned}$$

Our paper [29] used two-sided ideals as class sets of multiple classifiers. Since the multiplication does not commute, the family of one-sided ideals turns out to be substantially larger. Here we consider finitely generated right ideals as cluster sets of multiple clusterers. This will make it possible to investigate larger families of clusterers. Note that the situation of left ideals is dual to that of right ideals.

As customary, in order to avoid ambiguities that may occur in considering semigroup rings, zeros of semigroups are denoted by  $\theta$ , and the symbol  $0$  stands for the zero of a ring. As usual, the sets  $\{0\}$  and  $\{\theta\}$  will be also denoted by  $0$  and  $\theta$ , respectively.

Rees matrix semigroups and associated notions of completely  $0$ -simple semigroups and Rees quotients are very well known in semigroup theory and play crucial roles in describing the structure of semigroups. See [20–22, 34] for examples of recent results concerning these constructions.

Suppose that  $G$  is a group,  $I$  and  $\Lambda$  are nonempty sets, and  $e$  is the identity of  $G$ . As usual, we denote by  $G^1 = G \cup \{1\}$  and  $G^0 = G \cup \{\theta\}$  the group  $G$  with identity  $1$  and, respectively, zero  $\theta$  adjoined in a standard fashion. Let  $P = [p_{\lambda i}]$  be a  $(\Lambda \times I)$  matrix with entries  $p_{\lambda i} \in G^0$ , for all  $\lambda \in \Lambda, i \in I$ . The *Rees matrix semigroup*  $M^0(G; I, \Lambda; P)$  over  $G$  with *sandwich matrix*  $P$  consists of all triples  $(g; i, \lambda)$ , where  $i \in I, \lambda \in \Lambda$ , and  $g \in G^0$ , where all triples  $(\theta; i, \lambda)$  are identified with  $\theta$ , and where multiplication is defined by the rule

$$(g_1; i_1, \lambda_1)(g_2; i_2, \lambda_2) = (g_1 p_{\lambda_1 i_2} g_2; i_1, \lambda_2).$$

A *Brandt semigroup* is a Rees matrix semigroup with identity sandwich matrix (see [29] for references to recent results on Brandt semigroups).

If  $G$  is a group,  $M = M^0(G; I, \Lambda; P)$ , and  $i \in I, \lambda \in \Lambda$ , then we use standard notation for the sets

$$\begin{aligned} G_{*\lambda} &= \{(g; i, \lambda) \mid g \in G, i \in I\}, \\ G_{i*} &= \{(g; i, \lambda) \mid g \in G, \lambda \in \Lambda\}, \\ G_{i\lambda} &= \{(g; i, \lambda) \mid g \in G\}. \end{aligned}$$

Let  $S$  be a subsemigroup of the Rees matrix semigroup  $M^0(G; I, \Lambda; P)$ . The following standard notation will be used. For any  $i, \lambda \in I$ , set

$$\begin{aligned} S_{i\lambda} &= S \cap G_{i\lambda}, \\ S_{*\lambda} &= S \cap G_{*\lambda}, \\ S_{i*} &= S \cap G_{i*}. \end{aligned}$$

Further, for any subsets  $X \subseteq I, Y \subseteq \Lambda$ , let

$$\begin{aligned} S_{X*} &= \bigcup_{i \in X} S_{i*}, \\ S_{*Y} &= \bigcup_{\lambda \in Y} S_{*\lambda}. \end{aligned}$$

A semigroup is said to be *right (left) simple* if it has no proper *right (left) ideals*. A semigroup is *left (right) cancellative* if  $xy = xz$  ( $yx = zx$ ) implies  $y = z$ , for all  $x, y, z \in S$ . A semigroup is called a *right (left) group* if it is right (left) simple and left (right) cancellative.

We need a few known properties of right or left groups, as collected in the next lemma. A *band* is a semigroup entirely consisting of idempotents. A band is called a *left zero (right zero) band* if it satisfies the identity  $xy = x$  ( $xy = y$ ).

**LEMMA 3.1** [10, Theorem 1.27]. *For any finite semigroup  $S$ , the following are equivalent:*

- (i)  $S$  is right (left) simple.
- (ii)  $S$  is a right (left) group.
- (iii)  $S$  is isomorphic to the direct product of a right (left) zero band and a group.
- (iv)  $S$  is a union of its left (right) ideals and each of these ideals is a group.

### 4. Main results

Suppose that  $T = M^0(G; I, \Lambda; P)$  is a Rees matrix semigroup with a subsemigroup  $S$ . Denote by  $L = L(S)$  the set

$$L = L(S) = \left\{ \lambda \in \Lambda \mid S_{*\lambda} \subseteq \theta \text{ or } \bigcup_{i \in I} p_{\lambda i} S_{i*} \subseteq \theta \right\},$$

where  $p_{\lambda i} S_{i*} = \{(p_{\lambda i} g; i, \mu) \mid (g; i, \mu) \in S_{i*}\}$ . In particular,  $p_{\lambda i} S_{i*} \subseteq \theta$  means that  $p_{\lambda i} = \theta$  or  $S_{i*} = \emptyset$ .

**THEOREM 4.1.** *Let  $T = M^0(G; I, \Lambda; P)$  be a Rees matrix semigroup over a group  $G$  with zero and sandwich matrix  $P$ . If every column of the sandwich matrix  $P$  has at most one nonzero entry, then for every finite subsemigroup  $S$  of  $T$ , the maximum number  $E_r$  of errors of binary clusterers, which can be corrected by a multiple clusterer defined in  $\mathbb{F}_0[S]$  by a cluster set of the form  $\text{id}_r(g_1, \dots, g_k)$ , is equal to*

$$E_r = \max \left\{ \left\lfloor \frac{|S_{*L}| - 1}{2} \right\rfloor, \max_{\lambda \in \Lambda \setminus L} \left\lfloor \frac{|S_{*\lambda}| - 1}{2} \right\rfloor \right\}, \tag{4.1}$$

where  $L = L(S)$ . Moreover, if  $|G| > 1$  and equality (4.1) holds for every finite subsemigroup  $S$  of  $T$ , then every column of the sandwich matrix  $P$  has at most one nonzero entry.

Our main theorem follows from a formula for the largest weights of right ideals  $\text{id}_r(g_1, \dots, g_k)$  in  $\mathbb{F}_0[S]$ , which we record as an auxiliary proposition here. Notice that  $2 \lfloor (x + 1)/2 \rfloor$  is equal to the smallest even integer which is not more than  $x$ , and  $2 \lfloor x/2 \rfloor$  is equal to the largest even integer which is less than  $x$ .

**PROPOSITION 4.2.** *With the notation of Theorem 4.1 the following conditions are equivalent.*

- (i) *For every finite subsemigroup  $S$  of  $T$  and  $L = L(S)$ , the largest weight  $W_r$  of the cluster sets of the form  $\text{id}_r(g_1, \dots, g_k)$  in  $\mathbb{F}_0[S]$  is equal to*

$$W_r = \max \left\{ |S_{*L}|, \max_{\lambda \in \Lambda \setminus L} |S_{*\lambda}| \right\}. \tag{4.2}$$

- (ii) *Every column of the sandwich matrix  $P$  has at most one nonzero entry.*

Inverse semigroups form an important class and have been investigated by many authors. As an illustration here we cite only a few articles [2, 13–15, 17, 19, 31] and refer the readers to our previous paper [29] for a broader bibliography. It is well known that a Rees matrix semigroup over a group with zero is inverse if and only if every row and column of the sandwich matrix contains precisely one nonzero entry. Therefore Theorem 4.1 and Proposition 4.2 apply to inverse Rees matrix semigroups.

The following small example shows that it is impossible to replace two implications of Theorem 4.1 by two equivalent conditions as in Proposition 4.2.

**EXAMPLE 1.** Let  $T = M^0(G; I, \Lambda; P)$ , where  $G = \{e\}$ ,  $I = \{i\}$ ,  $\Lambda = \{\mu, \nu\}$ , and  $P = \begin{bmatrix} e \\ e \end{bmatrix}$ . Then the only column of  $P$  has two nonzero entries. Nevertheless, a tedious but routine verification shows that equality (4.1) holds for all subsemigroups  $S$  of  $T$ , because all clusters of the form indicated in Theorem 4.1 cannot correct any errors of the binary clusters.

### 5. Proofs

The left annihilator  $\text{Ann}_\ell(S)$  of  $S$  is the set defined by

$$\text{Ann}_\ell(S) = \{x \in S \mid x \neq \theta, xS = \theta\}.$$

It follows from the definitions of a Rees matrix semigroup and the sets  $L$  and  $S_{*L}$  that

$$\text{Ann}_\ell(S) = S_{*L}. \tag{5.1}$$

**PROOF OF PROPOSITION 4.2.** (i) *implies* (ii). Suppose to the contrary that condition (i) holds, but there exist two nonzero entries in one of the columns of the sandwich matrix  $P$ . Denote these entries by  $p_{\mu j}$  and  $p_{\nu j}$ , where  $j \in I$ ,  $\mu, \nu \in \Lambda$ , and  $p_{\mu j}, p_{\nu j} \in G$ . Consider the elements

$$a = (p_{\mu j}^{-1}; j, \mu), b = (p_{\nu j}^{-1}; j, \nu) \in M^0(G; I, \Lambda; P).$$

It follows from the definition of a Rees matrix semigroup that  $a = a^2 = ba$  and  $b = b^2 = ab$ . Thus  $B = \{a, b\}$  is a right zero band.

Letting  $S = B \cup \{\theta\}$  and  $g = a - b \in \mathbb{F}_0[S]$ , we consider the cluster set  $\text{id}_r(g)$ . It is straightforward to verify that the weight  $W_r = \text{wt}(\text{id}_r(g)) = \text{wt}(g)$  in the left-hand side of (4.2) is equal to 2. However,  $S_{*L} = \emptyset$  implies  $|S_{*L}| = 0$ . Furthermore,

$$|S_{*\lambda}| = \begin{cases} 1 & \text{if } \lambda \in \{\mu, \nu\} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore the right-hand side of equality (4.2) is equal to 1. This contradicts condition (i) and completes the proof that (i) implies (ii).

(ii) *implies* (i). Suppose that every column of the sandwich matrix  $P$  has at most one nonzero entry. We prove that condition (i) holds.

Take an arbitrary subsemigroup  $S$  of  $M^0(G; I, \Lambda; P)$ . First, following the advice of a referee, let us handle the easy situation where  $S$  does not contain  $\theta$ . Consider two elements  $(g_1; i_1, \lambda_1)$  and  $(g_2; i_2, \lambda_2)$  of  $S$  (possibly identical). Any product of these

in either order is nonzero. However, no column of the sandwich matrix  $P$  contains more than one nonzero entry by (ii). Hence  $\lambda_1 = \lambda_2$ . Putting  $\mu = \lambda_1$ , we see that all matrix entries  $p_{\mu i_1}$  and  $p_{\mu i_2}$ , which occur in the products we have just looked at above, are nonzero. It follows that every nonempty subset  $S_{i\mu}$  is a subsemigroup of  $S$  and is contained in a subgroup  $T_{i\mu}$  of  $T$ . Every finite subsemigroup of a group is also a group. Obviously, every  $S_{i\mu}$  is a right ideal of  $S$  and  $S$  is a union of these right ideals. It follows from Lemma 3.1 that  $S$  is a left group. Therefore  $L = \Lambda \setminus \{\mu\}$ ,  $\text{Ann}_\ell(S) = \emptyset$ ,  $S_*L = \emptyset$  and  $|S_*L| = 0$ . Furthermore,  $\Lambda \setminus L = \{\mu\}$  implies that  $S = S_\mu$ . Therefore in this situation equality (4.2) simplifies to  $W_r = |S|$ .

The inequality  $W_r \leq |S|$  is obvious. On the other hand, consider the element  $g = \sum_{s \in S} s$ . It follows from Lemma 3.1 that  $Ss = S$  for every  $s \in S$ . Hence  $g$  generates a one-dimensional right ideal  $\text{id}_r(g) = \mathbb{F}g$  in  $\mathbb{F}_0[S]$ . Therefore  $\text{wt}(\text{id}_r(g)) = \text{wt}(g) = |S|$ . Thus, condition (ii) is satisfied if  $S$  does not contain  $\theta$ .

In the rest of the proof we assume that  $\theta \in S$ . Choose a right ideal  $C = \text{id}_r(g_1, \dots, g_k)$  with the largest weight  $W_r$  among all weights of right ideals of this form in  $\mathbb{F}_0[S]$ . Obviously, all right ideals  $\text{id}_r(g_1), \dots, \text{id}_r(g_k)$  are contained in  $C$ . It is clear that the weight of a nonzero subideal is never less than the weight of an ideal containing it. Hence, by the maximality of  $W_r$ , all right ideals  $\text{id}_r(g_1), \dots, \text{id}_r(g_k)$  have the same weight  $W_r$  too. Therefore it is enough to prove condition (i) in the case where  $k = 1$ . Hence we further assume that  $C$  is generated by one element  $g = g_1$ , so that  $C = \text{id}_r(g)$  and  $W_r = \text{wt}(C)$ .

Denote the maximum in the right-hand side of equality (4.2) by  $R$ . We prove two inequalities  $W_r \geq R$  and  $W_r \leq R$ . This is divided into two parts below.

*Part 1.* First, we verify that

$$W_r \geq R. \tag{5.2}$$

Let us choose and fix an element  $\lambda \in \Lambda$  such that

$$|S_{*\lambda}| = \max_{\lambda \in \Lambda \setminus L} |S_{*\lambda}|. \tag{5.3}$$

Consider two possible cases.

*Case 1.*  $|S_{*L}| \geq |S_{*\lambda}|$ . Then  $R = |S_{*L}|$  and we have to prove that  $W_r \geq |S_{*L}|$ . To this end it suffices to demonstrate that  $\mathbb{F}_0[S]$  contains an element  $g_L$  generating a right ideal with

$$\text{wt}(\text{id}_r(g_L)) = |S_{*L}|. \tag{5.4}$$

Let us consider the element

$$g_L = \sum_{x \in S_{*L}} x \in \mathbb{F}_0[S]. \tag{5.5}$$

We claim that  $g_L$  satisfies (5.4).

First of all, it is clear that the weight of the element  $g_L$  itself is equal to

$$\text{wt}(g_L) = |S_{*L}|. \tag{5.6}$$

Now consider the whole right ideal  $\text{id}_r(g_L)$ . Take any  $s \in S$ . By (5.1), we get  $S_{*L} \cdot s = \theta$ . Hence  $g_L \cdot s = 0$  in  $\mathbb{F}_0[S]$  by the definition of a contracted semigroup ring. By the definition of a right ideal, we see that the right ideal  $\text{id}_r(g_L)$  is equal to the linear space  $\mathbb{F}g_L$  spanned by  $g_L$  in  $\mathbb{F}_0[S]$ . Therefore the weight of the whole right ideal  $\text{id}_r(g_L)$  is equal to the weight of its generator  $g_L$ . Thus (5.6) yields (5.4), as required.

*Case 2.*  $|S_{*L}| < |S_{*\lambda}|$ . Then (5.3) implies that  $|S_{*\lambda}| = \max_{\gamma \in \Lambda} \{|S_{*\gamma}|\} = \max_{\gamma \in \Lambda \setminus L} \{|S_{*\gamma}|\}$ . Hence  $R = |S_{*\lambda}|$  and we have to prove that  $W_r \geq |S_{*\lambda}|$ . By the maximality of  $W_r$ , to this end it suffices to find an element  $g_\lambda \in \mathbb{F}_0[S]$  which generates a right ideal with

$$\text{wt}(\text{id}_r(g_\lambda)) = |S_{*\lambda}|. \tag{5.7}$$

Notice that  $|S_{*\lambda}| > 0$ , because  $|S_{*\lambda}| > |S_{*L}|$ . Consider the element

$$g_\lambda = \sum_{x \in S_{*\lambda}} x. \tag{5.8}$$

We claim that  $g_\lambda$  satisfies (5.7).

Pick a nonzero element  $x$  with minimal weight in  $\text{id}_r(g_\lambda)$ . Since  $\text{wt}(x) = \text{wt}(\text{id}_r(g_\lambda))$ , it suffices to show that  $\text{wt}(x) = |S_{*\lambda}|$ . The proof of this claim is nontrivial and relies on the maximality in the choice of  $\lambda$ . After the proof we include Example 2, which demonstrates that the maximality of  $|S_{*\lambda}|$  is indeed essential in this step of our proof.

By the definition of a right ideal, there exists  $r \in \mathbb{F}_0[S] \cup \{1\}$  such that  $x = g_\lambda r$ . If  $r = 1$ , then  $x = g_\lambda$  and the claim follows. Hence we may further assume that  $r \in \mathbb{F}_0[S]$ . Since  $\mathbb{F}$  has only two elements, the definition of a semigroup algebra implies that

$$x = \sum_{s \in \text{supp}(r)} g_\lambda \cdot s. \tag{5.9}$$

Consider any element  $s \in \text{supp}(r)$ . There exist  $g \in G, i \in I, \mu \in \Lambda$  such that  $s = (g; i, \mu)$ . If  $p_{\lambda i} = 0$ , then  $S_{*\lambda} \cdot s = \theta$  implies that  $g_\lambda \cdot s = 0$ , and so we can subtract  $s$  from  $r$  and remove  $s$  from  $\text{supp}(r)$  to simplify (5.9). Thus, we may further assume that  $p_{\lambda i} \neq 0$ , and so  $p_{\lambda i} \in G$ , for all  $s = (g; i, \mu) \in \text{supp}(r)$ .

Now, for any  $s \in (g; i, \mu) \in \text{supp}(r)$ , using the definition of a Rees matrix semigroup (5.9), the assumption  $p_{\lambda i} \in G$  and the fact that  $G$  is a group, it is straightforward to verify that  $|S_{*\lambda}s| = |S_{*\lambda}|$  and  $S_{*\lambda}s \subseteq S_{*\mu}$ . The maximality of  $|S_{*\lambda}|$  shows that  $|S_{*\lambda}s| = |S_{*\mu}|$ , and so  $S_{*\lambda}s = S_{*\mu}$ . Thus we see that the following equalities hold:

$$g_\lambda s = \sum_{t \in S_{*\mu}} t, \tag{5.10}$$

$$|S_{*\mu}| = |S_{*\lambda}|. \tag{5.11}$$

If the number of elements  $s \in \text{supp}(r) \cap S_{*\mu}$  is even, then it follows from what we have proved above that the sum of all products  $g_\lambda s$ , for all of these elements  $s$ , is

equal to zero, because  $\mathbb{F}$  has characteristic 2. If, however, the number of the elements  $s \in \text{supp}(r) \cap S_{*\mu}$  is odd, then it follows that the sum of all products  $g_\lambda s$ , for all of these elements  $s$ , is equal to  $\sum_{t \in S_{*\mu}} t$ .

Therefore it follows from (5.9), (5.10) and (5.11) that there exist elements  $\mu_1, \dots, \mu_\ell$  in  $\Lambda$  such that

$$\begin{aligned} x &= \sum_{t \in S_{*\mu_1}} t + \dots + \sum_{t \in S_{*\mu_\ell}} t, \\ |S_{*\lambda}| &= |S_{*\mu_1}| = \dots = |S_{*\mu_\ell}|, \\ \text{supp}(x) &= S_{*\mu_1} \cup \dots \cup S_{*\mu_\ell}, \end{aligned}$$

where each of the sums  $\sum_{t \in S_{*\mu_1}} t, \dots, \sum_{t \in S_{*\mu_\ell}} t$  is obtained by collecting similar terms from (5.9), and so all of these sums belong to the right ideal  $\text{id}_r(g_\lambda)$ .

By the minimality of the weight of  $x$  in the right ideal  $\text{id}_r(g_\lambda)$ , we get  $\ell = 1$ . Therefore the weight of the whole ideal  $\text{id}_r(g_\lambda)$  is equal to  $\text{wt}(x) = |S_{*\lambda}|$ . Thus we have found the desired element  $g_\lambda$  in Case 2, as required.

Since one of the two cases above always occurs, this completes the proof of inequality (5.2).

We now prove that the reversed inequality

$$\text{wt}(\text{id}_r(g)) \leq R \tag{5.12}$$

holds for all elements  $g \in \mathbb{F}_0[S]$ . Clearly, it suffices to consider only one element  $g$ , which is chosen in  $\mathbb{F}_0[S]$  with the property that the weight  $\text{wt}(\text{id}_r(g))$  achieves the largest possible value.

Pick a nonzero element  $x$  of minimal weight in  $\text{id}_r(g)$ . We have

$$\text{wt}(\text{id}_r(g)) = \text{wt}(x).$$

If  $\text{supp}(x) \subseteq S_{*L}$ , then we get  $\text{wt}(x) \leq |S_{*L}| \leq R$ , and so (5.12) is satisfied. Hence it remains to consider the case where  $\text{supp}(x)$  is not contained in  $S_{*L}$ .

Then there exists an element  $(h; i, \lambda)$  in  $\text{supp}(x) \setminus S_{*L}$ , where  $i \in I, \lambda \in \Lambda \setminus L$  and  $h \in G$ . Since  $\lambda \notin L$ , it follows from the definition of the set  $L$  that  $p_{\lambda j} S_{j*} \neq \theta$  for some  $j \in I$ . Hence we can choose  $y \in S_{j*}$  such that

$$p_{\lambda j} y \neq \theta. \tag{5.13}$$

There exists  $v \in \Lambda$  such that  $y \in S_{jv}$ . Since the  $j$ th column of the sandwich matrix  $P$  has at most one nonzero entry  $p_{\lambda j} \neq \theta$ , we see that

$$p_{\mu j} = \theta \quad \text{for all } \mu \neq \lambda. \tag{5.14}$$

Let us now regard  $y$  as an element of  $\mathbb{F}_0[S]$ . Since  $G$  is a group, it follows from the definition of a Rees matrix semigroup and (5.13) that the mapping  $z \mapsto zy$  from  $S_{*\lambda}$  to  $S_{*v}$  is injective. By (5.14),  $p_{\lambda j} \neq \theta$  implies that the product  $xy$  is nonzero.

Given that  $y$  is a homogeneous element of  $\mathbb{F}_0[S]$ , we see that  $|\text{supp}(xy)|$  does not exceed  $|\text{supp}(x)|$ . Clearly,  $xy$  also belongs to the right ideal generated by  $g$ . Therefore it follows from the minimality of the weight of  $x$  that  $|\text{supp}(x)| = |\text{supp}(xy)|$ .

Since  $y \in S_{j\nu}$ , we get  $\text{supp}(xy) \subseteq S_{*\nu}$  by the definition of a Rees matrix semigroup. Therefore  $|\text{supp}(x)| \leq |S_{*\nu}|$ . If  $\nu \in L$ , then  $|\text{supp}(g)| \leq |S_{*L}| \leq R$ . On the other hand, if  $\nu \notin L$ , then  $|S_{*\nu}| \leq R$  by the definition of  $R$ . Hence  $|\text{supp}(g)| \leq R$  again.

Thus, we see that (5.12) always holds. This completes the proof of Proposition 4.2.  $\square$

The following example shows that our proof cannot be simplified by omitting the choice of  $\lambda$  at the beginning of Part 1 and claiming that the weight of every right ideal generated by  $g_i = \sum_{x \in S_{*i}} x$  is greater than or equal to  $|S_{*i}|$  for each  $i$ . Notice that this is true for  $i = \lambda$ . Now we show that it may be wrong for some other values of  $i$ .

**EXAMPLE 2.** Let  $G = Z_4 = \{e, g, g^2, g^3\}$  be the cyclic group of order 4 with identity  $e$ , let  $P$  be the identity matrix over  $G$ , and let  $I = \Lambda = \{1, 2, 3\}$ . In the Rees matrix semigroup  $M^0(G; I, \Lambda; P)$  consider the subset

$$S = \{\theta, (e; 1, 2), (g; 1, 2), (g^2; 1, 2), (e; 1, 3), (g; 1, 3), (g^2; 1, 3), (g^3; 1, 3), (e; 2, 3), (g; 2, 3)\}.$$

It is routine to verify that  $S$  is a subsemigroup of  $M^0(G; I, \Lambda; P)$  and  $S^2$  is contained in  $S_{13}$ .

Here we have  $|S_{*2}| = |S_{12}| = 4$ , and so  $2\lfloor(|S_{*2}| + 1)/2\rfloor = 4 = |S_{*2}|$ . However, letting

$$g_2 = (e; 1, 2) + (g; 1, 2) + (g^2; 1, 2),$$

we get  $g_2 \cdot ((e; 2, 3) + (g; 2, 3)) = (e; 1, 3) + (g^3; 1, 3)$ . Therefore

$$\text{wt}(\text{id}_r(g_2)) = 2 < |S_{12}|.$$

**PROOF OF THEOREM 4.1.** For an arbitrary  $G$ , the first assertion of our theorem with equality (4.1) immediately follows from Proposition 4.2, because we know that every multiple clusterer with a cluster set of weight  $W_r$  can correct  $\lfloor(W_r - 1)/2\rfloor$  errors of binary clusterers.

To prove the second assertion, consider the case where  $|G| > 1$ . As in the proof of Proposition 4.2, suppose to the contrary that there are two nonzero entries  $p_{\mu j}$  and  $p_{\nu j}$  in one of the columns of the sandwich matrix  $P$ . Put

$$S = \{\theta\} \cup T_{j\mu} \cup T_{j\nu}$$

and consider the element

$$g = \sum_{h \in G} [(h; j, \mu) - (h; j, \nu)].$$

Let  $C = \text{id}_r(g)$ . Since the characteristic of  $\mathbb{F}$  is equal to 2, it follows from the definition of a Rees matrix semigroup that  $C = \mathbb{F}g$ . Since  $|G| > 1$ , we see that  $\text{wt}(C) = \text{wt}(g) = 2|G|$ . Therefore the combined clusterer with the cluster set  $C$  can correct  $|G| - 1$  errors of the binary clusterers. However,  $S_{*L} = \emptyset$  implies that  $|S_{*L}| = 0$ . Furthermore,  $|S_{*\lambda}| \leq |G|$  for all  $\lambda$ , a contradiction with (4.1), because  $|G| - 1 > \lfloor(|G| - 1)/2\rfloor$ . This completes the proof of Theorem 4.1.  $\square$

## 6. Corollaries

**REMARK.** The case of left ideals is dual to that of right ideals. Hence as immediate corollaries we get dual versions of Theorem 4.1 and Proposition 4.2 answering analogous questions for left ideals.

Our proof shows that previous results obtained in [29] can be recorded in the following equivalent form. First [29, Theorem 2 and Example 4] in combination are equivalent to the following.

**COROLLARY 6.1** [29]. *Let  $T = M^0(G; I, \Lambda; P)$  be a completely 0-simple semigroup over a group  $G$  with zero and sandwich matrix  $P$ . If  $T$  is an inverse semigroup, then for every finite subsemigroup  $S$  of  $T$ , the maximum number  $E$  of errors of binary clusterers, which can be corrected by a multiple clusterer with cluster set  $\text{id}(g_1, \dots, g_k)$  in  $\mathbb{F}_0[S]$ , is equal to*

$$E = \left\lfloor \frac{\max\{M_Z, M_X, M_Y, M_G\} - 1}{2} \right\rfloor, \quad (6.1)$$

where  $X = \{i \in I \mid S_{i*} = \theta\}$ ,  $Y = \{\lambda \in \Lambda \mid S_{*\lambda} = \theta\}$ , and

$$\begin{aligned} M_Z &= |S_{Y*} \cap S_{*X}|, \\ M_X &= \max\{|S_{i*} \cap S_{*X}| : i \notin X\}, \\ M_Y &= \max\{|S_{*\lambda} \cap S_{Y*}| : \lambda \notin Y\}, \\ M_G &= \max\{|S_{i\lambda}| : i \notin X, \lambda \notin Y\}. \end{aligned}$$

Furthermore, if  $|G| > 1$  and equality (6.1) holds for all finite  $S \subseteq T$ , then  $T$  is inverse.

The present paper deals with contracted semigroup rings. In this construction another formula obtained in [29] can be recorded more concisely as follows. Our proof shows that [29, Proposition 5 and Example 4] in combination are equivalent to the following.

**COROLLARY 6.2** [29]. *With the notation of Corollary 6.1 the following conditions are equivalent.*

- (i) *For every finite subsemigroup  $S$  of  $T = M^0(G; I, \Lambda; P)$ , the largest weight  $W$  of the cluster sets  $\text{id}(g_1, \dots, g_k)$  in  $\mathbb{F}_0[S]$  is equal to*

$$W = \max\{M_Z, M_X, M_Y, M_G\}. \quad (6.2)$$

- (ii)  *$T$  is inverse.*

## 7. Practical applications and open question

The design of multiple clusterers by combining individual binary clusterers is quite common in the literature. This method has been used in various application areas

by many researchers (see [36, Section 7.5]). For example, it can be applied to all data sets available online for multiple clustering tasks in the UCI Machine Learning Repository [3].

Following Witten and Frank [36, Section 4.6], here we briefly mention only one situation, where it is absolutely necessary to combine several binary clusterers. Suppose that support vector machines are to be used in clustering. These are quite efficient, but can produce only binary clusterings. This is why one always has to combine several support vector machines in one scheme in order to design a multiple clusterer.

Our paper makes the very first step in research of the properties of Rees matrix constructions essential for clustering of data. It contains only theoretical results of independent interest, which can be used to guide future experiments for determining how the constructions perform in various practical situations.

There do not exist exact and conclusive theoretical criteria of efficiency in classification and data mining. This is also confirmed by the so-called ‘no free lunch’ theorems in search, optimization, and machine learning (see, for example, [37]). The efficiency of applications is always decided on the basis of experimental research. The results of experiments usually depend on the particular application area and are evaluated using statistical methods (see, for example, [7]).

The present paper is motivated by applications and aims to obtain results essential for guiding the design of future experimental work. Theorem 4.1 gives a complete description of the multiple clusterers in this construction, which are optimal with respect to property (A1) in Section 2. This result is essential, because in the design of experiments researchers are first of all interested in the cluster sets satisfying the basic properties (A1)–(A4). However, the following natural question motivated by property (A2) remains open.

**QUESTION 1.** For each positive integer  $m$ , each Rees matrix semigroup  $T = M^0[G; I, \Lambda; P]$ , and every finite subsemigroup  $S$  of  $T$ , describe all cluster sets of the form  $\text{id}(g_1, \dots, g_k)$  and  $\text{id}_r(g_1, \dots, g_k)$  in  $\mathbb{F}_0[S]$  with weight  $m$  and the largest possible information rate.

An answer to this question could help in the choice of cluster sets satisfying the basic properties (A1) and (A2) simultaneously for experimental evaluation in the future. Several separate experimental publications by a number of authors would be required for conclusive evaluation of the practical performance of applications like this in various areas. We refer to the monographs [36, 40] and recent articles [23, 27, 33, 35, 38, 39] for examples of experimental investigations of this kind.

### Acknowledgements

The authors are grateful to two referees for comments and corrections which have helped to improve this paper. In particular, the present more general versions of the main theorem and proposition were suggested by one of the referees.

## References

- [1] I. M. Araújo, A. V. Kelarev and A. Solomon, 'An algorithm for commutative semigroup algebras which are principal ideal rings with identity', *Comm. Algebra* **32**(4) (2004), 1237–1254.
- [2] C. J. Ash, T. E. Hall and J.-E. Pin, 'On the varieties of languages associated with some varieties of finite monoids with commuting idempotents', *Inform. and Comput.* **86**(1) (1990), 32–42.
- [3] A. Asuncion and D. J. Newman, *UCI Machine Learning Repository* (University of California, School of Information and Computer Science, Irvine, CA, 2009), <http://www.ics.uci.edu/~mllearn/MLRepository.html>.
- [4] A. M. Bagirov, A. M. Rubinov and J. L. Yearwood, 'A global optimization approach to classification', *Optim. Eng.* **3** (2002), 129–155.
- [5] A. M. Bagirov and J. L. Yearwood, 'A new nonsmooth optimization algorithm for minimum sum-of-squares clustering problems', *European J. Oper. Res.* **170** (2006), 578–596.
- [6] L. M. Batten, R. S. Coulter and M. Henderson, 'Extending abelian groups to rings', *J. Aust. Math. Soc. Ser. A.* **82**(3) (2007), 297–314.
- [7] S. Boslaugh and P. Watters, *Statistics in a Nutshell: A Desktop Quick Reference* (O'Reilly & Associates, Sebastopol, CA, 2008).
- [8] J. Cazaran and A. V. Kelarev, 'Generators and weights of polynomial codes', *Arch. Math. (Basel)* **69** (1997), 479–486.
- [9] J. Cazaran, A. V. Kelarev, S. J. Quinn and D. Vertigan, 'An algorithm for computing the minimum distances of extensions of BCH codes embedded in semigroup rings', *Semigroup Forum* **73** (2006), 317–329.
- [10] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups* (American Mathematical Society, Providence, RI, 1961).
- [11] M. J. Crabb, C. M. McGregor and W. D. Munn, 'A property of the complex semigroup algebra of a free monoid', *J. Aust. Math. Soc. Ser. A.* **81**(1) (2006), 97–104.
- [12] B. A. Davey, M. Jackson, M. Maróti and R. N. McKenzie, 'Principal and syntactic congruences in congruence-distributive and congruence-permutable varieties', *J. Aust. Math. Soc.* **85**(1) (2008), 59–74.
- [13] D. Easdown, J. East and D. G. FitzGerald, 'A presentation of the dual symmetric inverse monoid', *Internat. J. Algebra Comput.* **18** (2008), 357–374.
- [14] D. Easdown and W. D. Munn, 'Trace functions on inverse semigroup algebras', *Bull. Aust. Math. Soc.* **52**(3) (1995), 359–372.
- [15] S. M. Goberstein, 'Inverse semigroups determined by their partial automorphism monoids', *J. Aust. Math. Soc. Ser. A.* **81**(2) (2006), 185–198.
- [16] T. E. Hall, 'The radical of the algebra of any finite semigroup over any field', *J. Aust. Math. Soc. Ser. A* **11** (1970), 350–352.
- [17] T. E. Hall, 'Biprefix codes, inverse semigroups and syntactic monoids of injective automata', *Theoret. Comput. Sci.* **32**(1–2) (1984), 201–213.
- [18] J. M. Howie, *Fundamentals of Semigroup Theory* (Clarendon Press, Oxford, 1995).
- [19] M. Jackson, 'The embeddability of ring and semigroup amalgams is undecidable', *J. Aust. Math. Soc. Ser. A* **69** (2000), 272–286.
- [20] M. Jackson, 'On the finite basis problem for finite Rees quotients of free monoids', *Acta Sci. Math. (Szeged)* **67** (2001), 121–159.
- [21] M. Jackson and M. Volkov, 'Undecidable problems for completely 0-simple semigroups', *J. Pure Applied Algebra*, to appear.

- [22] M. Kambites, 'The loop problem for Rees matrix semigroups', *Semigroup Forum* **76**(2) (2008), 204–216.
- [23] B. H. Kang, A. V. Kelarev, A. H. J. Sale and R. N. Williams, 'A new model for classifying DNA code inspired by neural networks and FSA', in: *Pacific Knowledge Acquisition Workshop, PKAW2006*, Guilin, China, Lecture Notes in Computer Science, 4303 (Springer, Berlin, Heidelberg, 2006), pp. 187–198.
- [24] A. V. Kelarev, *Ring Constructions and Applications* (World Scientific, River Edge, NJ, 2002).
- [25] A. V. Kelarev, *Graph Algebras and Automata* (Marcel Dekker, New York, 2003).
- [26] A. V. Kelarev, R. Göbel, K. M. Rangaswamy, P. Schultz and C. Vinsonhaler, *Abelian Groups, Rings and Modules*, Contemporary Mathematics, 273 (American Mathematical Society, Providence, RI, 2001).
- [27] A. Kelarev, B. Kang and D. Steane, 'Clustering algorithms for ITS sequence data with alignment metrics', in: *AI06: Advances in Artificial Intelligence, 19th Australian Joint Conference on Artificial Intelligence*, Lecture Notes in Computer Science, 4304 (Springer, Berlin, Heidelberg, 2006), pp. 1027–1031.
- [28] A. V. Kelarev and D. S. Passman, 'A description of incidence rings of group automata', *Contemp. Math.* **456** (2008), 27–33.
- [29] A. V. Kelarev, J. L. Yearwood and M. A. Mammadov, 'A formula for multiple classifiers in data mining based on Brandt semigroups', *Semigroup Forum* **78**(2) (2009), 293–309.
- [30] A. V. Kelarev, J. L. Yearwood and P. W. Vamplew, 'A polynomial ring construction for classification of data', *Bull. Aust. Math. Soc.* **79** (2009), 213–225.
- [31] M. V. Lawson, S. W. Margolis and B. Steinberg, 'Expansions of inverse semigroups', *J. Aust. Math. Soc. Ser. A.* **80**(2) (2006), 205–228.
- [32] S. R. López-Permouth, K. P. Shum and N. V. Sanh, 'Kasch modules and pV-rings', *Algebra Colloq.* **12**(2) (2005), 219–227.
- [33] S. McCombie, P. Watters, A. Ng and B. Watson, 'Forensic characteristics of phishing—petty theft or organized crime?', in: *Proc. 4th Internat. Conf. on Web Information Systems and Technologies, WEBIST* (INSTICC Press, Setúbal, Portugal, 2008).
- [34] N. R. Reilly, 'Varieties generated by completely 0-simple semigroups', *J. Aust. Math. Soc.* **84**(3) (2008), 375–403.
- [35] P. Watters, F. Martin and H. S. Stripf, 'Visual detection of LSB-encoded natural image steganography', *ACM Trans. Appl. Perception* **5**(1) (2008).
- [36] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* (Elsevier/Morgan Kaufmann, Amsterdam, 2005).
- [37] D. H. Wolpert, 'The lack of a priori distinctions between learning algorithms', *Neural Comput.* **8**(7) (1996), 1341–1390.
- [38] J. L. Yearwood, A. M. Bagirov and A. V. Kelarev, 'Optimization methods and the  $k$ -committees algorithm for clustering of sequence data', *Appl. Comput. Math.* **8**(1) (2009), 92–101.
- [39] J. L. Yearwood, B. H. Kang and A. V. Kelarev, 'Experimental investigation of classification algorithms for ITS dataset', in: *PKAW2008, Pacific Rim Knowledge Acquisition Workshop*, Hanoi, Vietnam, 15–16 December 2008 (part of PRICAI-08, Tenth Pacific Rim Internat. Conf. Artificial Intelligence), Lecture Notes in Computer Science, 5465 (Springer, Berlin, Heidelberg, 2009), pp. 262–272.
- [40] J. L. Yearwood and M. Mammadov, *Classification Technologies: Optimization Approaches to Short Text Categorization* (IGI Global, Hershey, PA, 2007).

A. V. KELAREV, Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia  
e-mail: [a.kelarev@ballarat.edu.au](mailto:a.kelarev@ballarat.edu.au)

P. WATTERS, Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia  
e-mail: [p.watters@ballarat.edu.au](mailto:p.watters@ballarat.edu.au)

J. L. YEARWOOD, Graduate School of Information Technology and Mathematical Sciences, University of Ballarat, PO Box 663, Ballarat, Victoria 3353, Australia  
e-mail: [j.yearwood@ballarat.edu.au](mailto:j.yearwood@ballarat.edu.au)